

*Keywords: image segmentation, interactive image segmentation, initialization fuzzy regions*

Wawan GUNAWAN [0000-0002-9034-459X]\*

# **FUZZY REGION MERGING WITH HIERARCHICAL CLUSTERING TO FIND OPTIMAL INITIALIZATION OF FUZZY REGION IN IMAGE SEGMENTATION**

## **Abstract**

*One of the most important goals in image segmentation is the process of separating the object parts from the image background. Image segmentation is also a fundamental stage in the development of other image applications such as object recognition, target tracking, computer vision, and biomedical image processing. Interactive image segmentation methods with additional user interaction are still popular in research. Interactive image segmentation aims to provide additional information through simple interactions, especially in images with complex objects. Interactive image segmentation with region merging processes has drawbacks, one of which is suboptimal region splitting due to soft color shades, blurred contours, and uneven lighting, referred to in this study as ambiguous regions. However, in the fuzzy region initialization stage after obtaining values from the marker process, there is a possibility of missing or suboptimal determination of fuzzy regions. This is because it only takes the highest gray level value for the background marker and the lowest gray level value for the object marker. In this study, fuzzy region merging using hierarchical clustering is proposed to find optimal initialization for fuzzy regions in image segmentation. Based on the experimental results, the proposed method can achieve optimal segmentation with an average misclassification error value of 2.62% for Natural Images and 9.33% for Dental Images.*

## **1. INTRODUCTION**

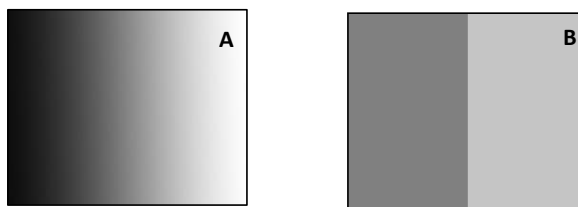
One of the most important goals in image segmentation is the process of separating the object and background parts of the image. Image segmentation is also a fundamental stage in the development of other image applications such as object recognition, target tracking, computer vision, and biomedical image processing. However, in specific images, especially biomedical ones (Makhlof et al., 2024), issues such as contours, color, lighting, and ambiguous regions make it difficult to produce optimal segmentation (Ning et al., 2010; Jung et al., 2014; Nguyen et al., 2013). Categorizing image segmentation into three categories: automatic segmentation, semi-automatic segmentation (interactive), and manual segmentation, for example using Adobe Photoshop application.

---

\* Universitas Islam Negeri Raden Intan Lampung, Indonesia, [wawan.gunawan@radenintan.ac.id](mailto:wawan.gunawan@radenintan.ac.id)

Currently, interactive image segmentation methods with user interaction remain popular and researched (Alemi Koohbanani et al., 2020; Mikhailov et al., 2024; Ding et al., 2023) with deep learning, (Militello et al., 2022; da Fonseca et al., 2021). Interactive image segmentation aims to provide additional information through simple interactions, especially for images with complex objects. In interactive image segmentation, user interaction greatly influences improving segmentation results. The better the information the user provides, the better the segmentation results (Jung et al., 2014). Generally, these methods are divided into 4 stages: First is region splitting, where the image is segmented into small regions according to its features. Second is the marker process to label the object and background areas, which can be lines, curves, etc. Third is feature extraction from each region, divided into 3 clusters: non-marker cluster, object marker cluster, and background marker cluster. Lastly, region merging combines small regions into segmented images.

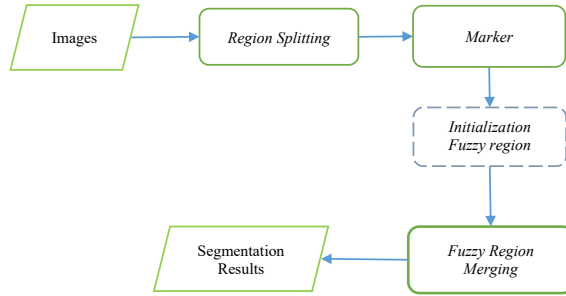
Interactive image segmentation with region merging has drawbacks, one of which is if the region splitting process is not optimal due to soft color shades, blurred contours, and uneven lighting, which are referred to as ambiguous regions in this study. The issue of ambiguous regions results in split regions potentially having two types of information, serving as both object and background. Hence, applying the Binary region Merging approach is not feasible as it would lead to over-segmentation. In Fig. 1 A, there is an example of an image with ambiguous regions where region splitting would be challenging, unlike in Fig. 1 B, where there is a clear separation for region splitting. In previous research, solving ambiguous region cases involved 4 stages: 1) Initial Segmentation, 2) Markers, 3) Fuzzy region initialization, 4) region merging using fuzzy similarity measurement (Gunawan et al., 2017). However, in the fuzzy region initialization stage, after obtaining values from the marker processes, missing or suboptimal fuzzy region determination is highly possible because it only takes the highest gray level value for the background marker and the lowest gray level value for the object marker. This research proposes fuzzy region merging using hierarchical clustering to find optimal initial fuzzy region initialization for image segmentation. The contribution of this research is the determination of optimal initial fuzzy regions using hierarchical clustering to find the most optimal fuzzy regions.



**Fig. 1. Difference in color transition in regions. (A) Ambiguous region, (B) Non-ambiguous region (Gunawan et al., 2017)**

## **2. RESEARCH METHOD**

In Figure 2, the steps carried out in the algorithm proposed by the authors can be seen, and the dashed boxes represent the areas that are the subject of this study and also serve as input to this study.



**Fig. 2. Proposed algorithm stages**

## 2.1. Region splitting

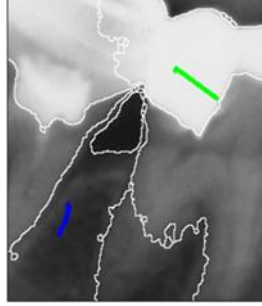
In the Region Splitting process, the authors employ the mean-shift algorithm using software from the Edison system (Edison Software, n.d.) to obtain initial segmentation. The mean-shift algorithm divides the image  $I$  into small regions  $I = [1, 2, \dots, r]$  based on the gradient of the probability density function in the image. The algorithm starts at the data point or pixel cluster center and then continuously shifts towards the highest pixel density until convergence is achieved for each cluster center. The final number of clusters depends greatly on the density level in the image. The mean-shift algorithm is quite effective in obtaining initial segmentation (Gunawan et al., 2017; Ning et al., 2010; Sankoh et al., 2016) because it considers both edges in the image and spatial information. Figure 3 shows the initial segmentation results using the mean-shift algorithm.

## 2.2. Marker

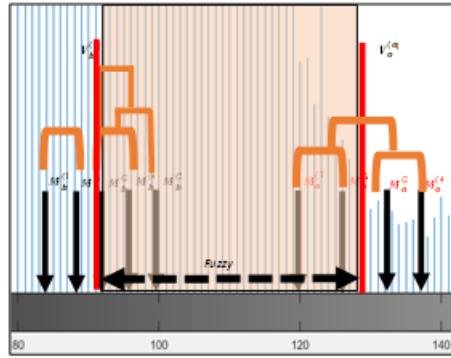
In interactive segmentation, users can provide additional information to assist the system in the segmentation process. In this study, user markings are used as a way to add information about object and background regions. In this research, the authors use user markings as a means to obtain information about objects and background. Figure 3 shows the marking process, where object regions are marked with green and background regions with blue. The marking process can be performed more than once depending on the complexity of the image. The determination of markers is highly sensitive, so it is necessary to search for regions with high similarity, not only in terms of the number of markers but also to ensure optimal fuzzy region determination (Sankoh et al., 2016).

## 2.3. Initialization of fuzzy regions

After the marking process is completed, it is time to extract information from the previous stages, namely region splitting and markers, to obtain fuzzy regions. Figure 4 illustrates the determination of fuzzy regions as done in one of the previous studies. The determination of fuzzy regions is highly sensitive because it affects the segmentation results. As we can see in Fig. 4, the fuzzy region area is between graylevel 100-120, if this shifts, the segmentation results will also change during the fuzzy region merging process. Furthermore, we can also conclude that the longer the fuzzy region, the higher the level of ambiguous regions in the image, and vice versa.



**Fig. 3. Results of region splitting and user marking process**



**Fig. 4. Determination of fuzzy regions is determined hierarchically through clustering**

The authors propose a different approach from previous methods, which is hierarchical clustering to find truly optimal fuzzy regions, which also serves as a contribution to this research. In Fig. 4, the authors also simulate the determination of fuzzy regions using hierarchical clustering approach. The calculation formula for fuzzy regions is inspired by previous methods (Arifin & Asano, 2005), where each marker result will be processed hierarchically, considering the variance mean of clusters. The formula we use is as follows:

$$m(C_k) = \frac{1}{P(C_k)} \sum_{z=T_{k-1}+1}^{T_k} zp(z) \quad (1)$$

Here are the definitions of the symbols in the formula:  $Am(C_k)$  is the mean of the cluster,  $P(C_k)$  is the function of the cluster  $C_k$  by calculating the frequency in the histogram for each gray level divided by the total pixels in the image ( $N$ ).  $T_k$  is the gray level in the region, and  $z$  is the occurrence frequency at gray level  $T_k$ , then  $p(z) = h(z)/N$ .

## 2.4. Fuzzy region merging

According to Pratamasunu et al. (2015) once the fuzzy regions are found, the final step is fuzzy region merging to obtain the segmented image result. The region merging process is performed on each region  $f_{i...r} \in F$ . The authors use fuzzy similarity measurement to calculate the initial seed region for the background  $C_B$  and the initial seed region for the object  $C_O$ . Fuzzy similarity measurement is calculated based on the similarity between graylevel and intensities, membership functions, and the difference in membership functions

towards their ordinal sets. The fuzzy similarity measurement  $\delta$  is calculated based on the global initial subset  $C_B$  and  $C_O$  against the local information in each fuzzy region  $f_i$  in the image. The similarity value  $\delta$  for the set  $(C_X U \{f_{ij}\})$ , initial seed  $C_X$ , membership function of gray levels in the fuzzy region  $f_i$ , and gray level intensity  $h(g)$  can be calculated using equation 2 (Gunawan et al., 2017).

$$(C_X U \{f_{ij}\}) = \frac{\sum_{g=1}^n (g - P(C_X U \{f_{ij}\}))}{\sum_{g=1}^n h(g)} \quad (2)$$

The fuzzy mean value  $P(A)$  represents the combined area  $A$ , calculated using the graylevel intensity  $h(g)$ , membership function  $\mu_A(g)$ , and the difference in membership functions towards their ordinal sets  $|(\mu_A(g) - \mu'_A(g))|$ , which can be calculated using equation 3. The result of fuzzy similarity measurement for each fuzzy region will be merged depending on the highest similarity value for each region, determining whether it will merge with the object or the background. The calculation of similarity to find the value of  $g$  in the fuzzy region  $\delta_{ig}$  can be calculated using equation 5.

$$P(A) = \sum_{g=1}^n h(g) \times \mu_A(g) \times |(\mu_A(g) - \mu'_A(g))| \quad (3)$$

$$\delta_{ig} = \operatorname{argmax}(\delta(C_B U \{f_{ij}\}) * (C_O U \{f_{ij}\})) \quad (4)$$

## 2.5. Test data

In evaluating the algorithm proposed by the authors, naturally datasets are required. The datasets used in this paper consist of two types: dental images obtained from the Airlangga University Hospital and natural images from the Weizmann's Segmentation Evaluation Database (Alpert et al., 2007).

## 2.6. Evaluation

Misclassification error (ME) is used to evaluate the segmentation results against the ground truth to assess segmentation performance. Equation 7 represents the formula for Misclassification error. Here,  $B_O$  and  $F_O$  represent the object and background pixels in the ground truth, while  $B_T$  and  $F_T$  represent the object and background pixels in the segmentation result.

$$ME = 1 - \frac{|B_O \cap B_T| + |F_O \cap F_T|}{|B_O| + |F_O|} \quad (5)$$

The authors also perform 2 other measurements to measure the performance of their proposed algorithm from edge detection:

1. Mean Squared Error (MSE): MSE is a measure of the average squared error between the edge detection results and the ground truth. The lower the MSE value, the better the quality of the edge detection results, because the error or deviation from the ground truth is smaller.
2. Peak Signal-to-Noise Ratio (PSNR): PSNR measures the quality of image reconstruction by comparing the edge detection results with the ground truth. The

higher the PSNR value, the better the quality of the edge detection results, because the noise relative to the signal (ground truth) is lower.

### 3. RESULTS AND DISCUSSION

In Table 1, the segmented image results for testing Natural and Panoramic Dental images are shown. The experiment was conducted on 5 Natural images and 5 Dental images, where all images were segmented well with Misclassification Error (ME) values. For Natural images, the smallest ME value was achieved in Test Image Number 4, which is 0.71%, while for Panoramic Dental images, it was in Image Number 6 with an ME value of 6.14. A smaller ME value indicates better segmentation results.

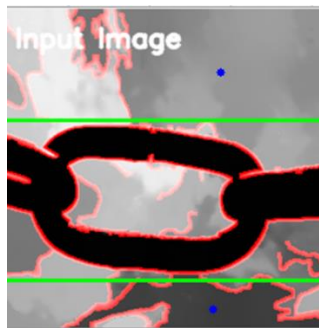


Fig. 5. Input Image on Graph-Cut with meanshift algorithm

The authors also compare the differences in the marker process in other algorithms, namely Graph-cut (Boykov & Jolly, 2001), so that the accuracy measurement is more balanced with their proposed algorithm. They also added a mean shift algorithm to get the initial marker and the addition of object and background marker information as shown in Figure 5.

Tab. 1. Segmentation result by calculating Misclassification error

No	Image Name	Grap-Cut with meanshift algorithm	Proposed Method
1	Natural 1	9.66%	2.32%
2	Natural 2	5.50%	3.85%
3	Natural 3	7.08%	1.5 %
4	Natural 4	2.32%	0.71 %
5	Natural 5	3.06%	4.7 %
6	Dental 1	7.35%	4.68%
7	Dental 2	37.5%	9.39 %
8	Dental 3	13.63%	12.14 %
9	Dental 4	14.14%	10.31%
10	Dental 5	30.03%	10.14%

Based on Table 1, the proposed method gets a value that tends to be consistently better than graph cut based on the ME value. The fuzzy approach to the region merging process

has proven to be quite effective in the segmentation process because it can overcome ambiguous regions, unlike graph cut which works binary.

The next test conducted was edge detection using the MSE and PSNR values as shown in Table 2. MSE in "Proposed" is lower than "Dahu Graph-cut": In general, the "Proposed" method has a lower MSE value for almost all images (both Natural and Dental groups), indicating that this method is more accurate and has a smaller error than the "Dahu Graph-cut" method. PSNR in "Proposed" is higher than "Dahu Graph-cut": The PSNR value of the "Proposed" method is generally higher than the "Dahu Graph-cut" method, indicating that the edge detection results of the "Proposed" method have better quality and are closer to the ground truth than the "Dahu Graph-cut" method.

**Tab. 2. Segmentation results by calculating Mean Squared Error and Peak Signal-to-Noise Ratio**

Image name	Proposed		Dahu Graph-Cut (Ôn Vũ Ngọc et al., 2023)	
	MSE	PSNR	MSE	PSNR
Natural 1	2481,8851	14,18	4407,4696	11,69
Natural 2	2128,6181	14,85	6476,6429	10,02
Natural 3	2778,8462	13,69	5881,2464	10,44
Natural 4	1946,7511	15,24	5831,3395	10,47
Natural 5	2417,0367	14,30	11803,5247	7,47
Dental 1	3133,0595	13,17	8045,5528	9,08
Dental 2	3524,4437	12,66	58844,9380	0,43
Dental 3	2675,5313	13,86	15982,0000	6,09
Dental 4	2537,2966	14,09	15572,0663	6,21
Dental 5	2059,5388	14,99	3473,7858	12,72

**Performance on Image Groups:**

- Natural Images: In Natural images, the "Proposed" method has a lower MSE and higher PSNR for all images compared to "Dahu Graph-cut." This indicates that the "Proposed" method performs better in detecting edges in natural images.
- Dental Images: In Dental images, the performance difference between the two methods is still visible, with the "Proposed" method showing lower MSE and higher PSNR. However, the difference between the two methods is more significant in some images such as Dental 2 and Dental 3, where the MSE for the "Dahu Graph-cut" method is much higher.

The Dahu Graph-cut method in the table shows less than optimal performance on medical images, especially on low contrast images, as seen in the Dental image group. This can be observed from the relatively high Mean Squared Error (MSE) value and the low Peak Signal-to-Noise Ratio (PSNR) in the Dahu Graph-cut method compared to the Proposed method.

The contribution of this research lies in hierarchical clustering to find the Optimal Initial fuzzy regions, where we search for the mean value in each cluster formed based on user markings. Table 3 shows the results of determining the mean value or hierarchical cluster points based on Gray Level in the images. The results of mean value A and mean value B will form fuzzy regions. For example, in Natural Image 2, the fuzzy region formed is in the gray level range between 27 to 159. These results will be processed for Region Merging using fuzzy similarity measurement. As shown in Table 1, Table 2 and Table 3, the method

we propose can perform segmentation well, although the determination of fuzzy regions still needs optimization by adding local region computation in calculating  $m(C_k)$ .

**Tab. 3. Mean cluster region values**

No	Image Name	Region cluster A (mean Value)	Region cluster B (mean Value)
1	Natural 1	16-41 (27)	38-236 (159)
2	Natural 2	2-52 (49)	192-239 (219)
3	Natural 3	2-2(2)	187-244(214)
4	Natural 4	6-21(12)	125-226(213)
5	Natural 5	5-25(19)	200-212(202)
6	Dental 1	79-96(85)	180-238(234)
7	Dental 2	23-24(24)	62-120(69)
8	Dental 3	39-42 (41)	212-239(217)
9	Dental 4	54-56(55)	101-212(112)
10	Dental 5	32-40(39)	181-234(224)

#### 4. CONCLUSION

The hierarchical clustering approach can be applied in determining fuzzy regions for the fuzzy region merging process. The proposed method is also capable of addressing ambiguous regions, resulting in more optimal segmentation. The segmentation results obtained an average ME value of 2.62% for Natural images and 9.33% for Dental images.

#### REFERENCES

- Alemi Koohbanani, N., Jahanifar, M., Zamani Tajadin, N., & Rajpoot, N. (2020). NuClick: A Deep Learning framework for interactive segmentation of microscopic images. *Medical Image Analysis*, 65, 101771. <https://doi.org/10.1016/j.media.2020.101771>
- Alpert, S., Galun, M., Basri, R., & Brandt, A. (2007). Image segmentation by probabilistic bottom-up aggregation and cue integration. *2007 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-8). IEEE. <https://doi.org/10.1109/CVPR.2007.383017>
- Arifin, A. Z., & Asano, A. (2005). Image thresholding by measuring the fuzzy sets. *Information Dan Technology Seminar* (pp. 189-194).
- Boykov, Y. Y., & Jolly, M.-P. (2001). Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images. *Eighth IEEE International Conference on Computer Vision. ICCV 2001* (pp. 105-112). IEEE. <https://doi.org/10.1109/ICCV.2001.937505>
- Da Fonseca, G. B., Perret, B., Negrel, R., Cousty, J., & Guimarães, S. J. F. (2021). Fuzzy-Marker-Based segmentation using hierarchies. In J. Lindblad, F. Malmberg, & N. Sladoje (Eds.), *Discrete Geometry and Mathematical Morphology* (Vol. 12708, pp. 391–403). Springer International Publishing. [https://doi.org/10.1007/978-3-030-76657-3\\_28](https://doi.org/10.1007/978-3-030-76657-3_28)
- Ding, Z., Wang, T., Sun, Q., & Chen, F. (2023). Rethinking click embedding for deep interactive image segmentation. *IEEE Transactions on Industrial Informatics*, 19(1), 261-273. <https://doi.org/10.1109/TII.2022.3157319>
- Gunawan, W., Arifin, A. Z., Indraswari, R., & Navastara, D. A. (2017). Fuzzy region merging using fuzzy similarity measurement on image segmentation. *International Journal of Electrical and Computer Engineering*, 7(6), 3402. <https://doi.org/10.11591/ijece.v7i6.pp3402-3410>
- Jung, C., Liu, J., Sun, T., Jiao, L., & Shen, Y. (2014). Automatic image segmentation using constraint learning and propagation. *Digital Signal Processing*, 24, 106-116. <https://doi.org/10.1016/j.dsp.2013.09.006>



Makhlouf, Z., Meraoumia, A., Lakhdar, L., & Haouam, M. Y. (2024). Enhancing medical data security in e-health systems using biometric-based watermarking. *Applied Computer Science*, 20(1), 28-55. <https://doi.org/10.35784/acs-2024-03>

Mikhailov, I., Chauveau, B., Bourdel, N., & Bartoli, A. (2024). A deep learning-based interactive medical image segmentation framework with sequential memory. *Computer Methods and Programs in Biomedicine*, 245, 108038. <https://doi.org/10.1016/j.cmpb.2024.108038>

Militello, C., Rundo, L., Dimarco, M., Orlando, A., Conti, V., Woitek, R., D'Angelo, I., Bartolotta, T. V., & Russo, G. (2022). Semi-automated and interactive segmentation of contrast-enhancing masses on breast DCE-MRI using spatial fuzzy clustering. *Biomedical Signal Processing and Control*, 71, 103113. <https://doi.org/10.1016/j.bsp.2021.103113>

Nguyen, T. N. A., Cai, J., Zheng, J., & Li, J. (2013). Interactive object segmentation from multi-view images. *Journal of Visual Communication and Image Representation*, 24(4), 477-485. <https://doi.org/10.1016/j.jvcir.2013.02.012>


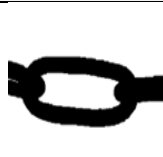

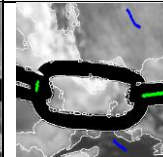
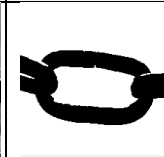



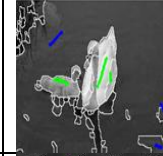


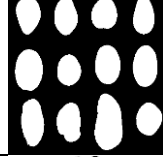
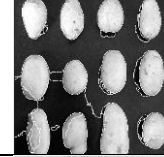
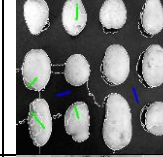
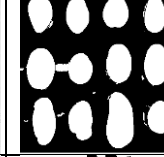



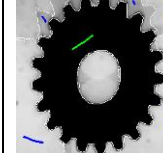

Ning, J., Zhang, L., Zhang, D., & Wu, C. (2010). Interactive image segmentation by maximal similarity based region merging. *Pattern Recognition*, 43(2), 445-456. <https://doi.org/10.1016/j.patcog.2009.03.004>

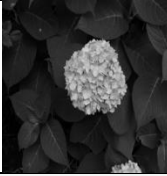
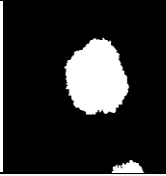
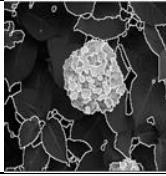
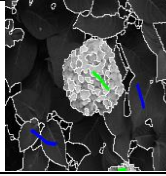

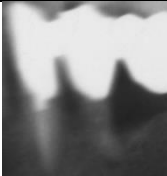


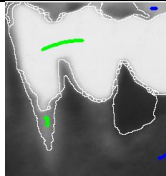



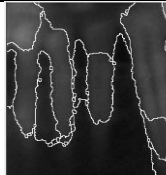
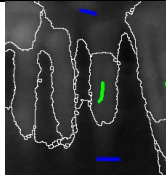

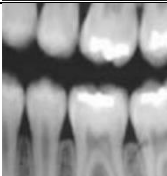

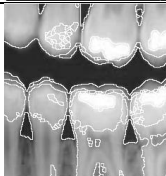
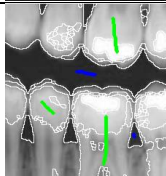



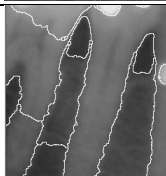
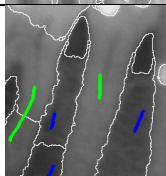

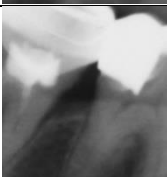

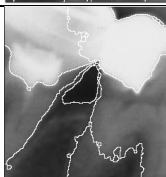
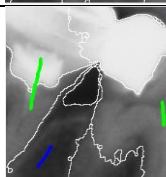

Ôn Vũ Ngọc, M., Carlinet, E., Fabrizio, J., & Géraud, T. (2023). The Dahu graph-cut for interactive segmentation on 2D/3D images. *Pattern Recognition*, 136, 109207. <https://doi.org/10.1016/j.patcog.2022.109207>

Sankoh, A. S., Arifin, A. Z., & Wijaya, A. Y. (2016). Extracted pixels similarity features (EPSF) using interactive image segmentation techniques. *International Journal of Computer Applications*, 136(2), 5-12. <https://doi.org/10.5120/ijca2016908236>

## APPENDIX

Tab. 1. Segmented image results

No	Original Image	Ground truth	Region Splitting	Marker	Segmentation Result	ME
1						2,32%
2						3,85%
3						1,5%
4						0,71%

5						4,7%
No	Original Image	Ground truth	Region Splitting	Marker	Segmentation Result	ME
6						4,68%
7						9,39%
8						12,14%
9						10,31%
10						10,14%